

Jussi Penttinmäki

SAN VOLUME CONTROLLERIN KÄYTTÖÖNOTTO

Opinnäytetyö
Kajaanin ammattikorkeakoulu
Luonnontieteiden ala
Tietojenkäsittelyn koulutusohjelma
Syksy 2014



Koulutusala Luonnontieteiden ala	Koulutusohjelma Tietojenkäsittelyn koulutusohjelman
Tekijä(t) Jussi Penttimäki	
Työn nimi San Volume Controllerin käyttöönotto	
Vaihtoehtoiset ammattiopinnot Järjestelmäylläpito	Toimeksiantaja Herman IT
Aika Syksy 2014	Sivumäärä ja liitteet 22+0
<p>Tämän työn tavoitteena on tutkia, miten IBM SAN Volume Controller asennetaan uuteen IT-ympäristöön ja ja kuinka palveluja siirretään jo olemassa olevista ympäristöistä tähän uuteen ympäristöön. Opinnäytetyö on tilannut Herman IT ja he ovat kiinnostuneita mahdollisuuksista siirtää palveluja datakeskuksien välillä käyttämällä IBM San Volume Controller.</p> <p>Käytännön työ tehdään Herman IT: n demoympäristössä, jossa käytännön työ tehdään ensin asentamalla IBM San Volume Controller IBM:n asentajan kanssa. Asennuksen jälkeen San Volume Controllerin suorituskykyä mitataan erilaisilla testi ohjelmistoilla.</p> <p>Opinnäytetyö teoriaosuus koostuu enimmäkseen IBM Redbookeista ja muusta käytettävissä olevasta materiaalista. Teoriaosuudessa vertaillaan San Volume Controlleria muihin vastaaviin tuotteisiin</p> <p>Opinnäytetyön tuloksena toimiva IT-ympäristö, jossa levyjärjestelmät on kytketty San Volume Controlleriin, josta levykapasiteetti jaetaan palvelimille.</p>	
Kieli	Suomi
Asiasanat	Levyjärjestelmävirtualisointi, Fibre Channel, iSCSI, SAN Volume Controller, SVC
Säilytyspaikka	<input checked="" type="checkbox"/> Verkkokirjasto Theseus <input type="checkbox"/> Kajaanin ammattikorkeakoulun kirjasto

School Business	Degree Programme Business Information Technology
Author(s) Jussi Penttimäki	
Title Implementing San Volume Controller	
Optional Professional Studies System Administration	Commissioned by Herman IT
Date Fall 2014	Total Number of Pages and Appendices 22+0
<p>Objective of this thesis is to study how to introduce IBM SAN Volume controller to new IT environment and migrating services from remote location to this new environment. Thesis is ordered by Herman IT and they are interested on the possibilities in datacenter migrations using IBM SAN Volume Controller.</p> <p>The practical work is made in the Herman IT's demo environment where the practical work is done by first installing IBM SVC with the help of IBM's technician. After installation of the SVC running storage benchmark programs to measure if the SVC makes difference in speed of the storage.</p> <p>Theory part of thesis is made mostly by studying IBM Redbooks and other available material. Also by comparing SVC to other similar products.</p> <p>The result of the thesis is working IT environment where storage systems are connected to host by SVC.</p>	
Language of Thesis Finnish	
Keywords	Storage virtualization,, Fibre Channel. iSCSI, SAN Volume Controller, SVC
Deposited at	<input checked="" type="checkbox"/> Electronic library Theseus <input type="checkbox"/> Library of Kajaani University of Applied Sciences

SISÄLLYS

1 JOHDANTO	1
2 LEVYJÄRJESTELMÄVIRTUALISOINTI	1
2.1 RAID-tasot	1
2.2 Storage Area Network	3
2.2.1 iSCSI	4
2.2.2 FC	5
2.2.3 FCoE	6
3 IBM SAN VOLUME CONTROLLER JA EMC VPLEX	7
3.1 IBM San Volume Controller	7
3.2 EMC VPLEX	11
4 IBM SAN VOLUME CONTROLLERIN ASENNUS JA KÄYTTÖÖNOTTO	14
4.1 Kaapelointi	15
4.2 Käynnistys ja alkukonfigurointi	15
4.3 San Volume Controllerin päivittäminen	16
4.4 Levyjärjestelmien lisääminen	17
4.5 Hostien lisääminen	17
4.6 Levyjakojen lisääminen	18
5 YHTEENVETO	21
LÄHTEET	22

SYMBOLILUETTELO

FC	Fibre Channel
FCoE	Fibre Channel over Ethernet
GUI	Graphical User Interface, eli Graafinen käyttöliittymä
Gb/s	Gigabittia per sekunti, mittayksikkö
Gt	Gigatavu, mittayksikkö
Host	Yleensä palvelin, jolle tarjotaan levykapasiteettia
LUN	Logical Unit Number, eli looginen levyn osa.
MDisk	Managed Disk Katso LUN
Ms	millisekuntti
RAID	Redundant array of independent disks tai redundant array of inexpensive disks
Release notes	Julkaisutiedot
Räkkikaappi	Erityisvalmisteinen kaappi, johon voidaan asentaa mm. palvelimia, levyjärjestelmiä ja kytkimiä
SAN	Storage Area Network
SVC	San Volume Controller
San Volume Controller	IBM:n levyjärjestelmä virtualisointi tuote

UPS	Uninterruptible power supply suojaa laitteita syöttöjännitteen epätasaisuuksilta ja virransyötön katkoksilta.
Unit	Yksikkö, jolla mitataan rakkikiinnitteisten laitteiden korkeutta.
VPLEX	EMC:n levyjärjestelmä virtualisointi tuote
Volume	Katso LUN
WWN	World Wide Name
WWNN	World Wide Node Name
WWPN	World Wide Port Name
iSCSI	Internet Small Computer System Interface

1 JOHDANTO

Tämä dokumentti kertoo levyjärjestelmävirtualisoinnista ja sen komponenteista, levyjärjestelmävirtualisointituotteista, sekä IBM San Volume Controllerin käyttöönotosta.

Olen valinnut tämän aiheen koska uskon, että sen tutkimisesta olisi minulle tulevaisuudessa hyötyä ja sen lisäksi aihe kiinnostaa minua. Olen myös ollut IBM:n järjestämällä Spektriakatemia kurssilla aiheesta.

2 LEVYJÄRJESTELMÄVIRTUALISOINTI

Levyjärjestelmävirtualisoinnilla tarkoitetaan usean fyysisen levyjärjestelmän yhdistämistä yhdeksi levylaitteeksi, jota hallitaan keskitetysti. Virtualisoinnin vuoksi levyjärjestelmien ylläpito selkeytyy ja helpottuu, koska et enää hallitse jopa monen eri valmistajan laitteita erikseen, vaan hallitset kaikkia yhdestä portaalista. (1)

Levyjärjestelmävirtualisointi toteutetaan storage hypervisorilla, jota myös kutsutaan software-defined storageksi. Storage hypervisorille voidaan lisätä eritasoista ja erivalmistajan tarjoamaa tallennus kapasiteettia, josta tallennus kapasiteetti jaetaan sitä käyttäville koneille. Eri tasoisella levyllä voidaan tarkoittaa esimerkiksi 7200 rpm SAS, 10k SAS tai SSD-levyjä. Näiden levyjen kapasiteetilla ja nopeudella on suuriakin eroja Gt per euro ja IOPS per euro mitattuna. (2)

Looginen levy tai looginen volume on levyosio joka voi olla osa fyysistä levyä tai voi koostua useasta fyysisestä levystä. Loogisia levyjä kutsutaan myös LUNeiksi. San Volume Controllerissa hosteille jaettavia LUNeja kutsutaan volumeiksi ja San Volume Controllerille jaettuja LUNeja kutsutaan MDiskeiksi, eli Managed Diskeiksi. (3)

I/O, eli Input/Output operaatio tarkoittaa kirjoitus ja lukuoperaatiota. IOps, eli Input/Output operations per second on tapa mitata kirjoitus ja lukuoperaatioiden määrää per sekunti. (4)

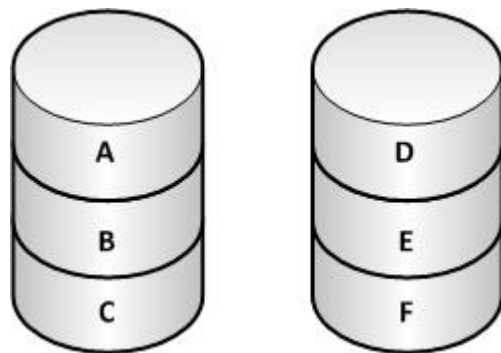
Levyjärjestelmät tallentavat datan blokki tasolla. Kun käyttäjä tallentaa tiedoston, jaetaan tiedosto ennalta sovitun kokoiisiin blokkeihin, jotka tallennetaan levyille. Kun käyttäjä haluaa avata tiedoston, koostetaan se levyllä olevista blokeista. Koska blokkikoko on ennalta sovittu ja sama koko levyllä, saattaa levyille jäädä tyhjää tilaa jota ei voida käyttää, koska tiedoston koko ei ole suoraan jaollinen levyn blokkikoolla. (5)

2.1 RAID-tasot

Yksinkertaisimmillaan storage virtualisointi on useasta levystä tehty RAID, jossa useasta fyysisestä levystä tehdään niin sanottu raidi-pakka. Tämä raidi pakka näkyy käyttöjärjestelmälle yhtenä levynä. Erilaisilla raidi pakkoilla saadaan datalle suojausta levyrikkoja vastaan, lisää ka-

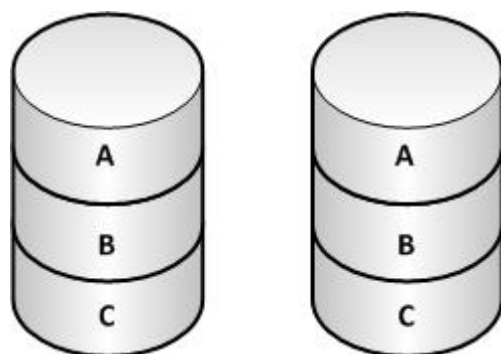
kapasiteettia. sekä kirjoitus ja lukunopeutta. Yleisimpiä raid ”tasoja” ovat raid 0, 1, 5, 6 ja 10.
(6)

Raid 0 pakka koostuu vähintään kahdesta levystä ja data kirjoitetaan tasaisesti levyille (Kuva 1). Tämän raidi tason hyviä puolia on kirjoitusnopeus ja kapasiteetti, koska data kirjoitetaan raidi pakkaan vain kerran. Huonoja puolia raid 0:ssa on suojauksen puuttuminen, koska data kirjoitetaan raidi pakkaan vain kerran, jo yhdenkin levyn hajoamin aiheuttaa datan menetyksen. (6)



Kuva 1: Raid 0

Raid 1 pakassa kahdesta levystä tehdään peilipari, jossa kaikki data kirjoitetaan molemmille levyille eli peilataan (Kuva 2). Raidi 1 pakan hyviä puolia on datan suojaus ja hyvä suorituskyky. Koska data on peilattu, voi toinen levy raidi pakasta rikkoutua ja dataa ei menetetä. Huonoja puolia raid 1 pakassa on ”hukkakapasiteetin” määrä. Puolet levykapasiteetistä kuluu datan suojaukseen. (6)

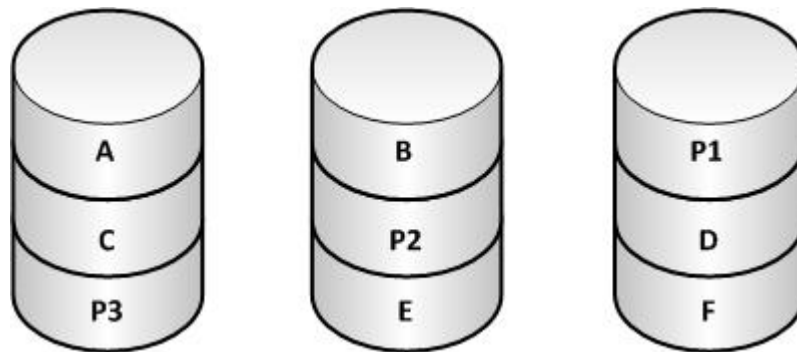


Kuva 2: Raid 1

Raid 10 on raid 0:n ja raid 1 yhdistelmä, jossa data kirjoitetaan aina kahdelle levyille. Tämä raidi taso on kaikkein kallein koska puolet kapasiteetistä kuluu datan suojaukseen. Raidi 10

kestää myös suurimman määrän levyrikkkoja, jos ”oikeat” levyt menevät rikki. Raidi 10 pakalla on myös hyvä suoritus kyky, koska data on jaettu eri levyjen kesken. (6)

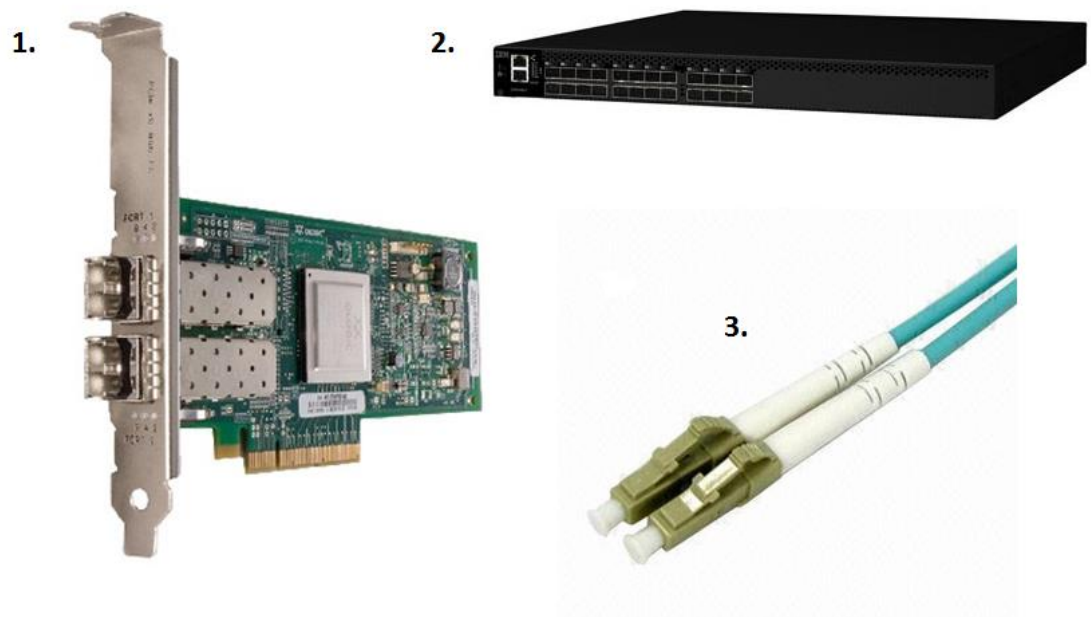
Raid 5 ja 6 pakoissa data kirjoitetaan tasaisesti usealle levyille, jonka lisäksi jokaiselle levyille kirjoitetaan pariteetti dataa (Kuva 3). Raid 5 ja 6 levyillä on hyvä lukunopeus, koska data sijaitsee usealla levyllä. Mutta myös huono kirjoitusnopeus, koska kirjoitetusta datasta joudutaan laskemaan pariteetti joka myös kirjoitetaan levyille. Raidi 5 ja 6 levyjen ero on pariteetti levyjen määrässä. Raid 5:ssä on yksi pariteetti levy ja raid 6:ssa on kaksi pariteetti levyä. Tästä johtuen raid 5 pakka kestää yhden levyn hajoamisen ja raid 6 kahden levyn hajoamisen. Pariteetti levyjen data on jaettu tasaisesti raidi pakkaan. Raidi 5 ja 6 pakat ovat hyviä vaihtoehtoja, jos halutaan kohtuullinen suojaus kohtuullisilla kuluilla. (6)



Kuva 3: Raid 5

2.2 Storage Area Network

Storage Area Network eli SAN, on tärkeä osa levyjärjestelmävirtualisointia. SAN on tallennuskapasiteetille dedikoitu verkko, jota hyväksikäyttäen keskitetystä tallennuskapasiteetista jaetaan kapasiteettia palvelimille esimerkiksi levyjen tai nauhureiden muodossa. SAN koostuu normaalisti kolmesta komponentista (kuva 4): host bus adaptereista (HBA) (1.), kytkimistä (2.) ja kaapeleista (3.). SAN -pohjaisissa laitteissa dataa käsitellään blokkitasolla, eikä sitä pidä sekoittaa NAS-pohjaisiin laitteisiin joissa dataa käsitellään tiedostotasolla. (7) (8)



Kuva 4: 1. Qlogic Fibre Channel host bus adapter (9) 2. IBM SAN24B-5 Fibre Channel kytkin (10) 3. Monimuotokuitukaapeli LC-LC liittimellä (11)

Levyjärjestelmävirtualisoinnissa käytetään kolmea eri SAN protokollaa. iSCSI, FC ja FCoE protokollaa.

2.2.1 iSCSI

iSCSI eli Internet Small Computer System Interface on blokkitasoinen TCP/IP pohjainen SAN protokolla. iSCSI käyttää tiedonsiirtoon 1Gb/s tai 10Gb/s TCP/IP verkkoja. Useita yhteyksiä voidaan kanavoida yhdeksi sessioksi kapasiteetin ja viansiedon parantamiseksi. Virheenkorjaukseen iSCSI käyttää TCP-protokollaa, joka uudelleen lähettää hukatut paketit. (12)

Koska iSCSI käyttää ethernet verkkoa, ei erillisiä kytkimiä tai verkkokortteja tarvita, vaan voidaan käyttää jo olemassa olevia laitteita. Tämä tekee iSCSI protokollan käyttöönotosta FC ja FCoE protokollia halvemman. Kuitenkin suositeltavaa on käyttää iSCSI:a omassa verkossa, koska iSCSI ei sisällä liikenteen salausta. iSCSI protokolla on erittäin tunnettu protokolla ja siihen liittyviä ongelmia voidaan ratkoa perinteisillä työkaluilla, kuten Wiresharkilla. (12)

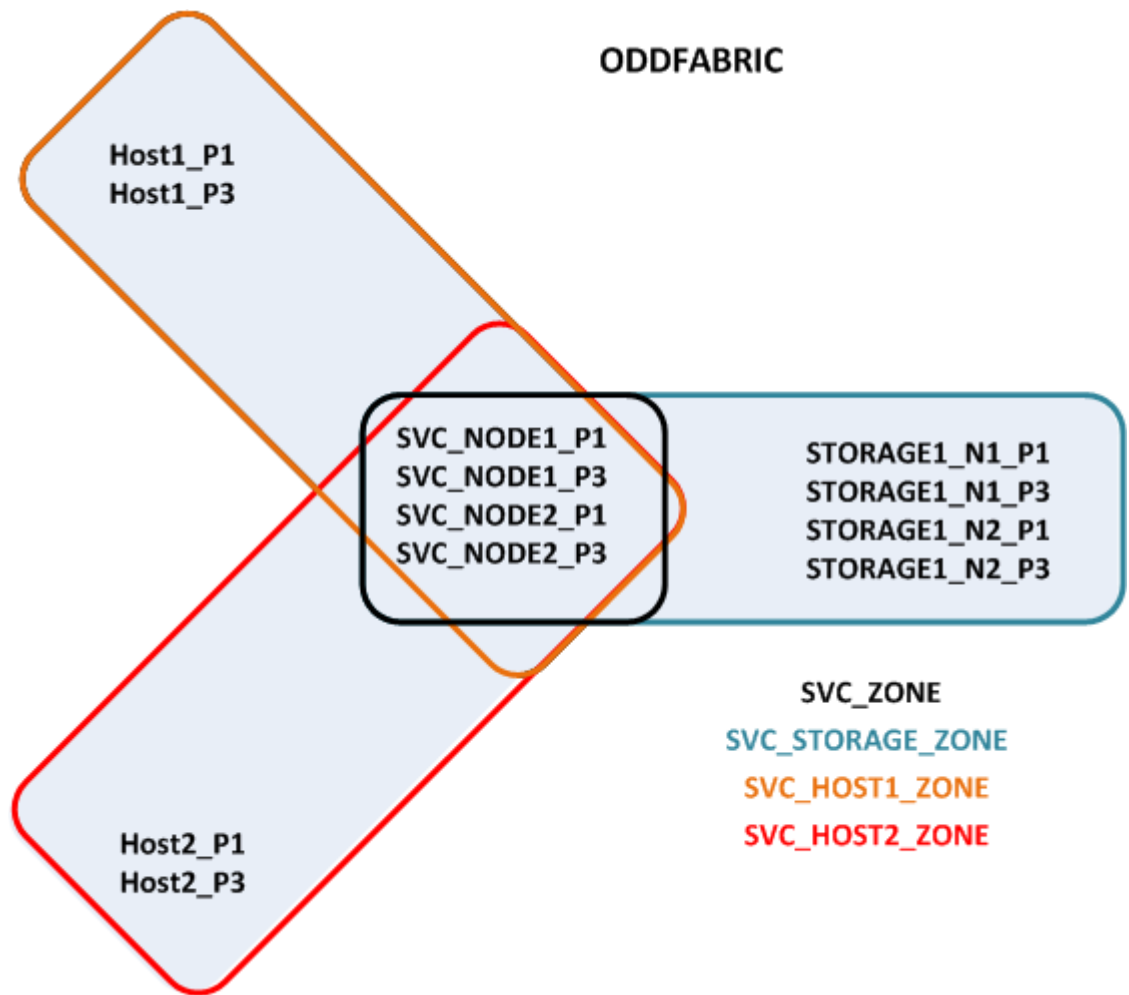
iSCSI initiaattoreita on kahdenlaisia, ohjelmistopohjaisia ja rautapohjaisia. Ohjelmistopohjaisessa iSCSI:ssa koko iSCSI pino toteutetaan ohjelmistolla, tästä johtuen ohjelmisto pohjainen iSCSI käyttää enemmän palvelimen prosessori tehoa ja muistia. Rautapohjaisessa iSCSI:ssa koko iSCSI pino toteutetaan erillisellä verkkokortilla, tästä johtuen rautapohjainen iSCSI käyttää vähemmän palvelimen prosessori tehoa ja muistia. (13)

2.2.2 FC

FC eli Fibre Channel on myös blokkitasen SAN protokolla. FC protokolla käyttää tiedonsiirtoon 1 Gb/s, 2 Gb/s, 4 Gb/s, 8 Gb/s tai 16 Gb/s Fibre Channel verkkoja. Fibre Channel verkoissa tallennuskapasiteettiin ollaan yhteydessä FC kehyksillä, joihin on kapseloitu SCSI komentoja ja dataa. Koska Fibre Channel on erittäin tunnettu ja luotettava protokolla käytetään sitä yleensä kriittisiin ympäristöihin. (12)

Fibre Channel verkon täytyy olla hävikitön, koska Fibre Channel protokollassa ei ole virheenkorjausta. Hävikitön verkko saadaan rajoittamalla liikennettä ruuhkan estämiseksi. Koska Fibre Channel ei käytä ethernet kytkimiä, vaan vaatii omat Fibre Channel kytkimet ja adapterit on se iSCSI:iin verrattuna erittäin kallis. Fibre Channel verkko on myös ethernet verkkoa vaikeampi konfiguroida. (12)

Fibre Channel verkko perustuu zoningiin. Fibre Channel verkossa eri laitteiden ja porttien tunnistamiseen käytetään WWN-osoitetta (Word Wide Name). WWN-osoite on ethernet verkon MAC-osoitetta vastaava osoite. WWN-osoite muodostuu 16 heksadesimaaliluvusta, jotka on jaettu 8:saan 2 luvun ryhmään. Ryhmät on erotettu toisistaan kaksoispisteellä. WWN-osoitteita on kahdenlaisia WWNN ja WWPN-osoitteita. WWNN eli World Wide Node Name käytetään tunnistamaan laite ja WWPN eli World Wide Port Name käytetään tunnistamaan portti. WWN-osoite esimerkki 80:05:07:68:01:30:df:01. WWN-osoitteille voidaan antaa alias helpottamaan konfigurointia. WWN-osoitteista tehdään erilaisia zoneja, joissa olevat portit pystyvät keskustelemaan keskenään.. Kuvassa 5 esimerkki toisen kytkimen zoneista ympäristöstä jossa on 2 kytkintä, 2 palvelinta, 1 levyjärjestelmä ja 2 San Volume Controller noodia. Kytkimen nimi on ODDFABRIC ja siihen on liitetty kummankin palvelimen adaptereista portit 1 ja 3, kummankin San Volume Controller noodin portit 1 ja 3, sekä levyjärjestelmän kummankin noodin portit 1 ja 3. (12)



Kuva 5: ODDFABRIC Zonet

2.2.3 FCoE

FCoE eli Fibre Channel over Ethernet on myös blokkitason protokolla. FCoE protokollassa SCSI komennot ja data kapseloidaan ethernet kehyksiin. FCoE protokollassa on paljon FC protokollan ominaisuuksia. FCoE on iSCSI ja Fibre Channel protokollaan verrattaessa erittäin tuore protokolla. Tästä ja konfiguroinnin monimutkaisuudesta johtuen ongelmien ratkominen saattaa olla haasteellista. (12)

FCoE vaatii 10Gb hävikittömän ethernet verkon, mikä lisää kustannuksia iSCS:iin verrattaessa. Samassa verkossa voidaan myös ajaa muuta liikennettä. FC protokollan tapaan FCoE:ssa tehdään zoneja liikenteen ohjaamiseksi. Koska FCoE framet ovat 2200 tavun kokoisia, eikä niitä voida jakaa, täytyy kytkimien tukea jumbo frameja. (12)

3 IBM SAN VOLUME CONTROLLER JA EMC VPLEX

IBM San Volume Controller ja EMC VPLEX ovat hyvin samankaltaisia levyjärjestelmävirtualisointilaitteita. Kumpikin tuote otetaan käyttöön pareittain, jolloin laite parit muodostavat toisiaan suojaavia klustereita. San Volume Controllerissa näitä klustereita kutsutaan I/O Groupeiksi ja VPLEXissä näitä klustereita kutsutaan Engineiksi. (14)

Vaikka kumpikin tuote on hyvin samankaltaisia, on tuotteissa kuitenkin eroja. Suurimpana erona on välimuistin toiminta. (14)

3.1 IBM San Volume Controller

San Volume Controller on IBM:n levyjärjestelmävirtualisointilaite, joka sijoitetaan levyjärjestelmien ja palvelimien väliin. San Volume Controllerilla saavutetaan tallennuskapasiteetin keskitetty hallinta, vikasietoisuuden parantuminen, porrastettu tallennuskapasiteetti, levyjärjestelmätason snapshotti, disaster recoveryn, sekä muita hyödyllisiä toimintoja.

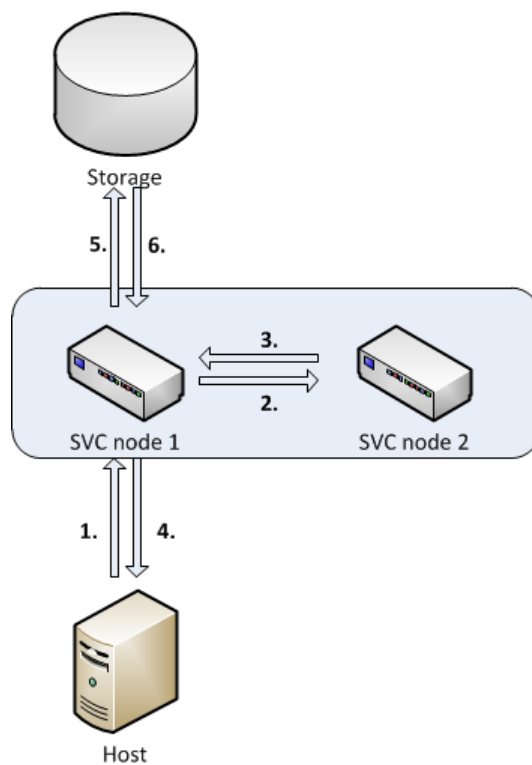
San Volume Controller on kehitetty San Josessa Kaliforniassa IBM: Almaden tutkimuslaitoksessa. Ohjelmisto laitteeseen on kehitetty Iso-Britanniassa Hursley laboratoriossa. Ohjelmisto on vuosian varrella kokenut suuria muutoksia etenkin käyttöliittymän ohella. Nykyisin San Volume Controllerin ohjelmistoa käyttää myös IBM Storwize tuoteperheen levyjärjestelmät. (15)

San Volume Controller noodin on IBM:n 1 unitin korkuinen muokattu rakkipalvelin, johon kuuluu 1 unitin korkuinen UPS. Noodi sisältää 24 Gt muistia ja 4 kappaletta 8Gb/s fibre channel kuituportteja. Noodeihin voidaan lisätä 24 Gt lisämuistia ja 4 kappaletta 8Gb/s fibre channel portteja tai 4 kappaletta 10Gb/s iSCSI portteja. 24 Gt muistin lisäys vaatii myös toisen prosessorin lisäämisen noodiin. San Volume Controllerin UPS:lla varmistetaan, ettei sähkökatkoksen aikana menetetä välimuistiin kirjoitettua dataa, jota ei ole vielä kirjoitettu loppupään varmistuskapasiteettiin. (16)

San Volume Controller klusteri koostuu 2 – 8 noodista, jotka muodostavat I/O ryhmiä. Jokainen I/O ryhmä koostuu kahdesta noodista, jotka tarjoavat vikasietoisuutta toisilleen. Jokaisella noodilla on 24 Gt välimuisti, joka peilataan I/O ryhmän välillä. (17)

San Volume Controllerille jaetaan loppupään tallennuskapasiteetista levyjä, taikka loogisia levyjä joita kutsutaan San Volume Controllerissa managed diskeiksi tai MDiskeiksi. MDiskit jaetaan yleensä levyjärjestelmistä FC tai FCoE protokollilla, jossa kiinto- tai SSD-levyistä on tehty erilaisia RAID-pakkoja. 1- 128 MDiskistä muodostetaan storage pooleja. MDisk voi olla vain yhden storage poolin jäsen. Storage poolia voidaan kasvattaa tai pienentää lisäämällä tai poistamalla MDiskejä storage poolista ilman käyttökatkoja. (17)

San Volume Controllerin kirjoitusoperaatio kuvan 6 mukaisesti. Palvelin suorittaa kirjoitusoperaation kyseistä MDiskia hallitsevalle noodille (1), jonka jälkeen noodit replikoi operaation saman I/O ryhmän toiselle noodille (2 ja 3), tästä johtuen San Volume Controllerin väli-
muistiin kirjoitettua dataa ei menetetä vaikka toinen I/O ryhmän noodit menetettäisiin. Tämän jälkeen hallitseva noodit kuittaa palvelimelle operaation suoritetuksi (4). Lopuksi hallitseva noodit suorittaa kirjoituksen lopulliseen taikka lopullisiin levyjärjestelmiin (5 ja 6). (17)

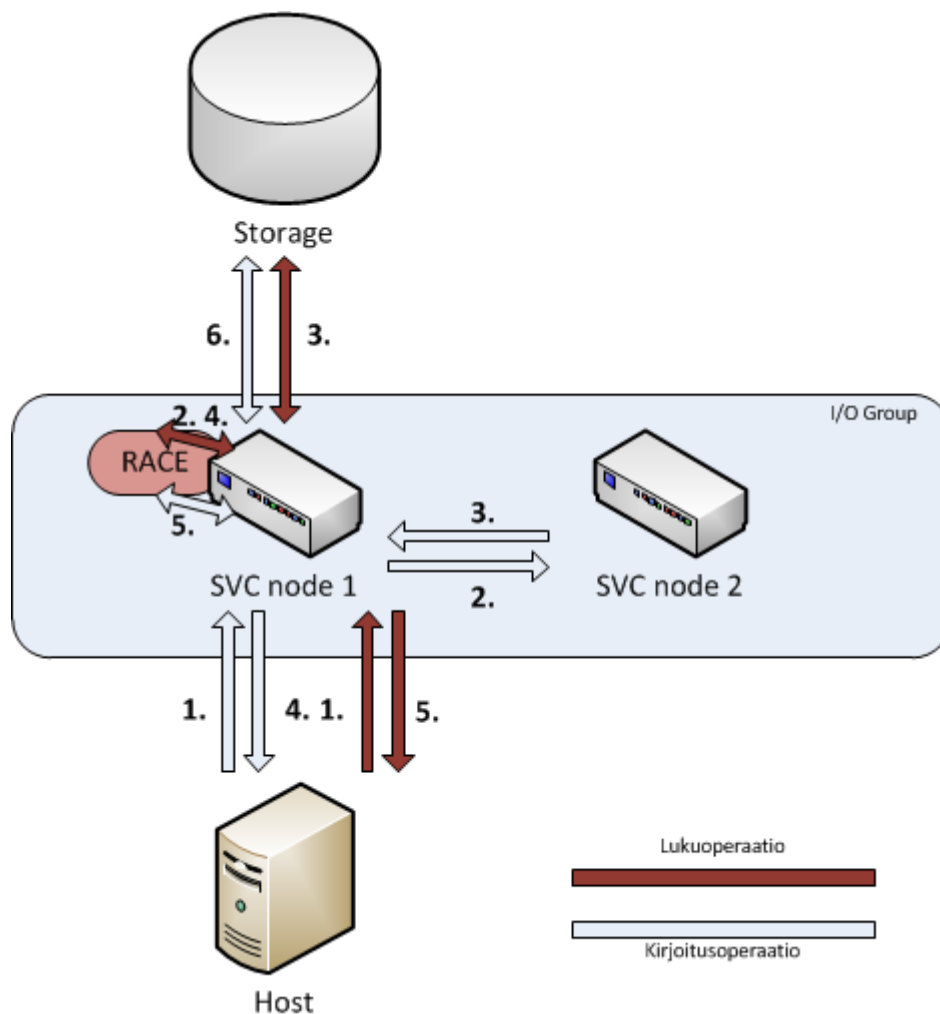


Kuva 6: San Volume Controller kirjoitusoperaatio

Easy Tier on San Volume Controllerin ominaisuus, jolla samassa storage poolissa voidaan käyttää monen eri tasoista levyä. SSD-levyt ovat huomattavasti tehokkaampia kuin kiintolevyt, mutta kiintolevyjen kapasiteetti on huomattavasti suurempi kuin SSD levyjen. Tällöin

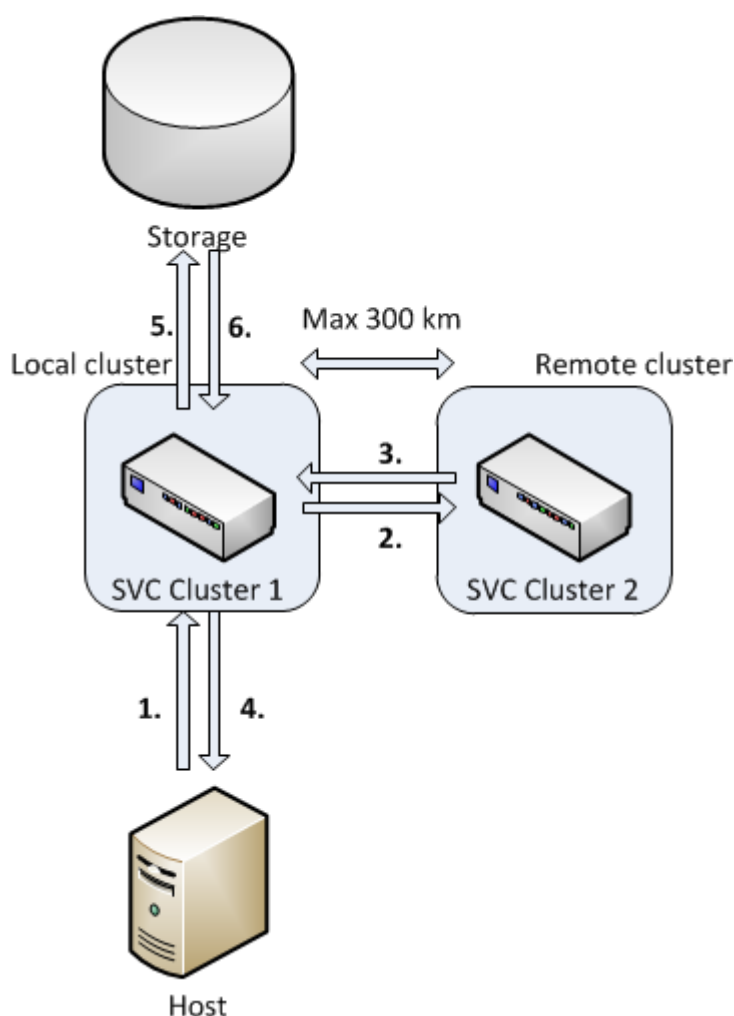
Easy Tier ominaisuutta käyttämällä saadaan säästöjä suurien datamäärien tallennuksessa, jossa vain osalla datalla on suuria kirjoitus- ja lukunopeus vaatimuksia. San Volume Controller automaattisesti siirtää datan jota käytetään paljon nopeammalle levyille, kun taas data jota käytetään vähemmän siirretään hitaammalle levyille. (17)

IBM Real-time Compression ja thin provision ovat ominaisuuksia, joilla reaaliaikaisesti säästetään tallennuskapasiteettia. Thin provisioidussa volumessa koko volumen kokoa ei varata storage poolista, kuten normaalissa volumessa, vaan vain käytetty data kirjoitetaan storage pooliin. IBM Real-time Compressionilla pakatussa volumessa, thin provisioinnin lisäksi, data pakataan San Volume Controlleriin sisäänrakennetulla RACE moottorilla kuva 7 mukaisesti. RACE moottori pakkaa ja purkaa datan reaaliaikaisesti. Kirjoitusoperaatiossa pakkaus tapahtuu San Volume Controller noodin välimuistiinkirjoittamisen ja lopulliseen tallennuskapasiteettiin kirjoittamisen välissä. Lukuoperaatiossa, jos dataa ei löydy San Volume Controllerin välimuistista, pyytää San Volume Controller RACE engineä lukemaan ja purkamaan lopullisesta tallennuskapasiteetista datan. (17)



Kuva 7: Real-time compression kirjoitus- ja lukuoperaatiot

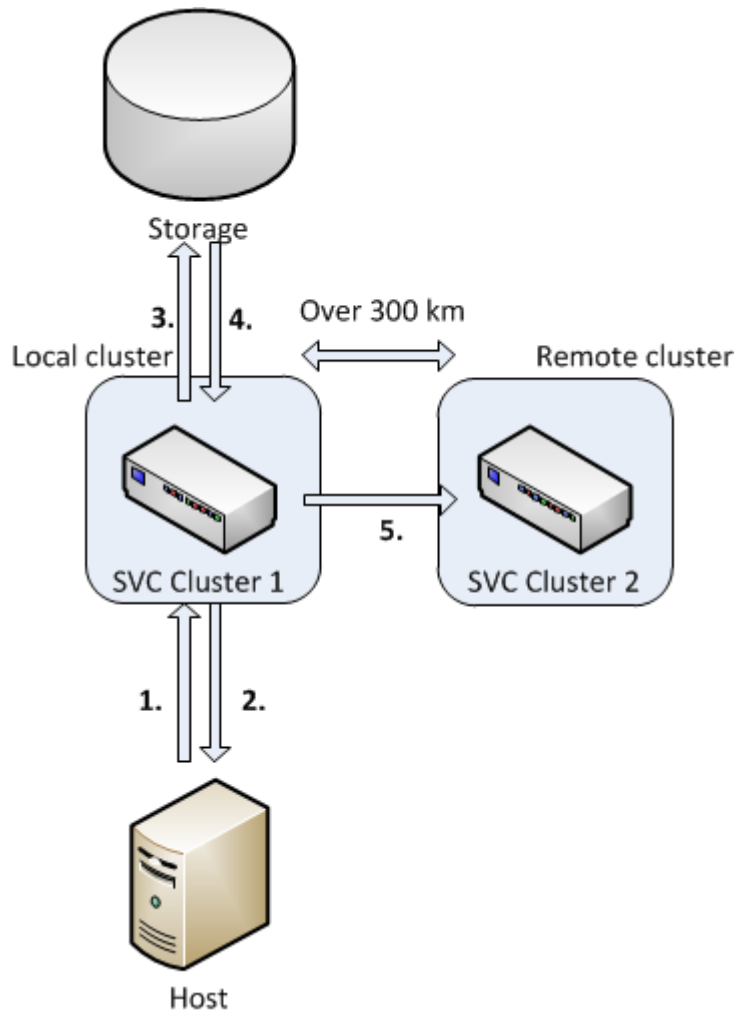
Metro ja Global mirror ovat San Volume Controllerin ominaisuuksia, joilla dataa voidaan peilata San Volume Controller klustereiden välillä. Metro mirror on täysin synkroninen peilaustapa, jossa kirjoitus operaatiossa paikallisen San Volume Controller noodin välimuistiin kirjoituksen ja kirjoitusoperaation palvelimelle kuitaamisen välissä operaatio synkronoidaan etä- San Volume Controller klusterin välimuistiin (Kuva 8). Tästä johtuen metro mirror klusterit voivat sijaita maksimissaan 300 km säteellä toisistaan. Metro Mirror klusterien etäisyys toisistaan vaikuttaa kirjoitusnopeuksiin ja vasteaikoihin negatiivisesti. (17)



Kuva 8: San Volume Controller Metro Mirror

Global mirror on asynkroninen peilaustapa, jolla San Volume Controller klusterit peilaavat dataa kahden San Volume Controller klusterin kesken kuten metro mirroringissa, mutta kirjoitusoperaatio kuitataan kirjoittavalle palvelimelle heti paikallisen San Volume Controller klusterin välimuistiin kirjoituksen jälkeen ja data peilataan etäklusteriin tämän jälkeen (Kuva

9). Global mirror peilauksessa ei San Volume Controller klustereiden välimatkalla ole maksimirajaa, eikä hidas ja suuren vasteajan omaava yhteys hidasta itse palvelimen toimintaa, toisin kuin metro mirroringissa. Kummallakin ominaisuudella voidaan toteuttaa disaster recovery site, jossa kriittisten palvelimien datat ovat onnettomuuden sattuessa heti käytettävissä, eikä aikaa vievää palautusta tarvitse tehdä. (17)



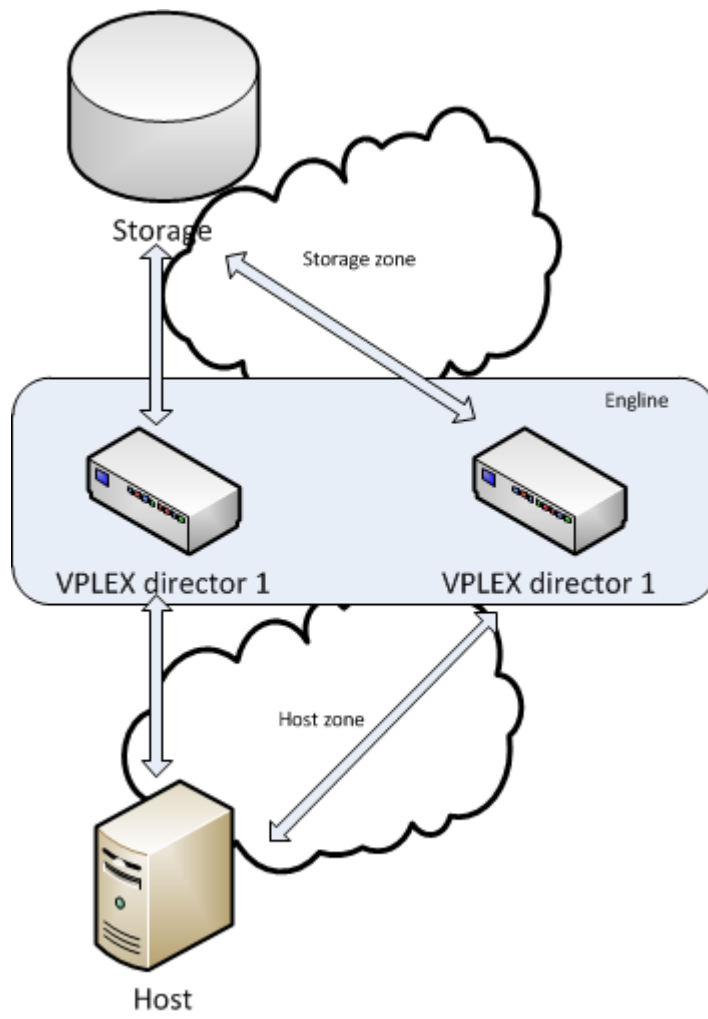
Kuva 9. San Volume Controller Global Mirror

3.2 EMC VPLEX

EMC:n VPLEX on storage virtualisointi laite, joka toimii levyjärjestelmien ja palvelimien välissä, kuten IBM San Volume Controller. VPLEX virtualisoi siihen liitettyjä levyjärjestelmiä, piilottamalla tallennuskapasiteetin oikean lähteen itse palvelimilta ja esittämällä sen omana tallennuskapasiteettina. (18)

VPLEX:stä on seuraavia hyötyjä: Koska VPLEX sijoittuu levyjärjestelmien ja palvelimien väliin, voidaan vplexiä käyttää datan siirtoon levyjärjestelmältä toiselle ilman käyttökatkoja palveluun. VPLEXin välimuisti toimii vain lukutoiminnoille eroten San Volume Controllerin välimuistista, joka toimii myös lukutoimintojen lisäksi myös kirjoitustoiminnoille. Koska kirjoitustoiminnot eivät käytä välimuistia, vaan kirjoitustoiminnot viedään noodin läpi suoraan levyjärjestelmään, ei VPLEX:n välimuistista ole samankaltaista hyötyä kuin San Volume Controllerin välimuistista. (18)

VPLEX voidaan ottaa käyttöön kolmella eri tapaa (Local, Metro ja Geo). VPLEX asennetaan aina pareittain. Local vaihtoehdossa VPLEX asennetaan paikallisesti yhteen konesaliin, jossa sillä hallitaan keskitetysti useita levyjärjestelmiä (Kuva 10). VPLEX Metrossa VPLEX asennetaan kahteen suhteellisen lähellä toisistaan oleviin konesaleihin, siten että konesalien välinen viive on maksimissaan noin 5 ms, jolloin VPLEXiä hyväksikäyttäen voidaan tallennuskapasiteettia käyttää kummastakin konesalista. On hyvä ottaa huomioon Metro mirrorin tuoma viive suunniteltaessa VPLEX Metro ympäristöä. VPLEX Geo asennetaan Metron tapaan kahteen konesaliin, mutta Geo:ssa konesalien etäisyys saa olla maksimissaan niin suuri, että tiedonsiirrossa on maksimissaan 50 ms viive. Metrossa tiedonsiirto on synkronista ja Geossa asynkronista. Kaikki VPLEX asennustavat käsittelevät dataa blokkitasolla. (19)



Kuva 10: EMC Local

4 IBM SAN VOLUME CONTROLLERIN ASENNUS JA KÄYTTÖÖNOTTO

Asennettava ympäristö koostuu San Volume Controlleristä, IBM DS levyjärjestelmästä, kahdesta IBM:n SAN kytkimestä, kahdesta FC palvelimesta, sekä yhdestä iSCSI palvelimesta. Palvelimet on myös kytketty erilliseen ethernet verkkoon. Kaikki laitteet asennetaan laboratorioympäristössä samaan rakkikaappiin, lukuun ottamatta iSCSI-yhteydellä olevaa palvelinta. San Volume Controller koostuu kahdesta 1Unit kokoisesta noodista, joilla olemmilla on oma 1U kokoinen UPS, sekä oma 1U kokoinen virranotin.

Asennus aloitetaan kiinnittämällä virranottimien kiinnitysraudat virranottoimeen, sekä kiinnittämällä UPSien ja San Volume Controller noodien kiskot rakkikaappiin. UPSien asennus paikaksi valittiin rakkikaapin alareuna, unitit 3 ja 4. Virranottimille unitit 5 ja 6, sekä itse San Volume Controller noodeilla unitit 7 ja 8, kuvan 11 mukaisesti.

8	San Volume Controller	
7	San Volume Controller	
6	Virranotin	
5	Virranotin	
4	UPS	
3	UPS	
2		
1		

Kuva 11: San Volume Controller sijoitus rakkikaappiin

Virranottimiin kiinnitetään virtakaapelit ja vedonpoistot. Tämän jälkeen virranotin kiinnitetään sille varattuun unittiin rakkikaapin takapuolelle. San Volume Controller noodit ja UPSit asennetaan kiskoihin. San Volume Controller noodeihin asennetaan myös kaapelien hallinta käsivarret.

Levyjärjestelmä, palvelimet ja kytkimet asennetaan San Volume Controllerien yläpuolelle.

4.1 Kaapelointi

Kaapelointi aloitetaan kytkemällä UPSien hallintakaapelit ja virtakaapelit UPSeista San Volume Controllereihin. Seuraavaksi kytketään verkkokaapelit kummankin San Volume Controllerin ethernet porttiin 1 ja ethernet kytkimen välille.

Kummankin San Volume Controllerin lisäkorttiin ja san-kytkimiin asennetaan 4 kappaletta 8Gb/s Fibre Channel optiikoita. Kuitukaapelointi suoritetaan käyttämällä multimode-kuituja taulukon 1 mukaisesti.

San Volume Controller noodit ja portit	SAN-kytkin ja portti
Noodi 1 Portti 1	SANSW1 portti 1
Noodi 1 Portti 2	SANSW2 portti 1
Noodi 1 Portti 3	SANSW1 portti 2
Noodi 1 Portti 4	SANSW2 portti 2
Noodi 2 Portti 1	SANSW1 portti 3
Noodi 2 Portti 2	SANSW2 portti 3
Noodi 2 Portti 3	SANSW1 portti 4
Noodi 2 Portti 4	SANSW2 portti 4

Taulukko 1: SAN kaapelointi

Kaikki San Volume Controllerin kaapelit asennetaan kaapelien hallintakouruun.

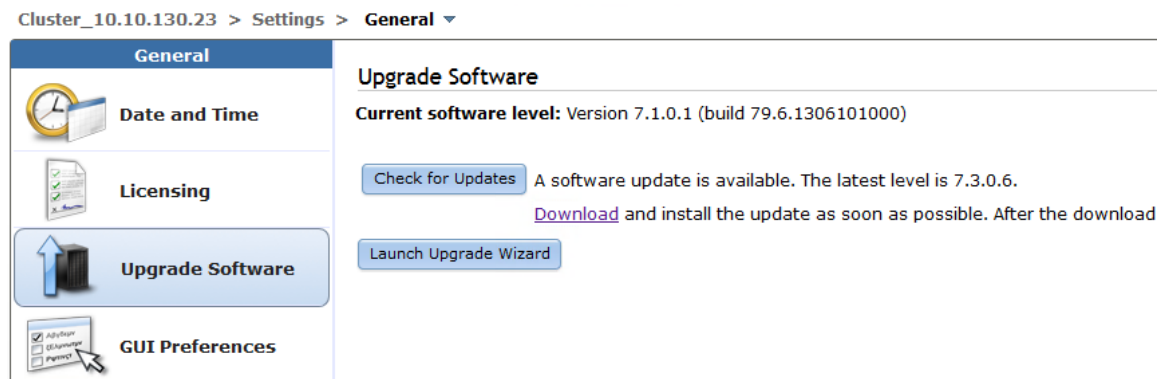
4.2 Käynnistys ja alkukonfigurointi

Ennen San Volume Controllerin käynnistystä UPSien etupaneeli täytyy irrottaa ja UPSien akun kytkennät täytyy kiinnittää. Tämän jälkeen molempien UPSien virtakaapelit voidaan

kytkää ja toinen San Volume Controller käynnistää. Käynnistämisen jälkeen San Volume Controller suorittaa erilaisia testejä, jotka kestävät muutamia minuutteja. Kun San Volume Controller on käynnistynyt, voidaan sille syöttää hallinta ip. Tämä tapahtuu noodin etupaneelin kautta. Kun ip osoite on syötetty, voidaan toinen noodi käynnistää, koska San Volume Controller toimii klusterissa, ei toiselle noodille tarvitse hallinta ip:tä, vaan liittyy jo olemassa olevaan klusteriin. Tämän jälkeen voidaan San Volume Controller klusterin konfigurointi aloittaa.

4.3 San Volume Controllerin päivittäminen

Klusterin päivittäminen tapahtuu San Volume Controllerin GUI:n kautta Settings > General > Upgrade Software sivulta (Kuva 12). IBM:n internet sivulta löytyy päivitystiedosto, test utility, sekä päivityksen release note tiedosto. Sivulta löytyy myös taulukko päivityspoluista. Release note tiedosto sisältää listan korjatuista ongelmista, uusista ominaisuuksista, vaatimuksista ja ohjeet päivitykseen. Test utility ajetaan ennen varsinaista päivitystä ja se testaa voidaanko päivitys asentaa klusteriin.



Kuva 12: San Volume Controller Upgrade Software

Päivityspaketti ladataan San Volume Controlleriin GUI:n kautta ja valitaan automatic upgrade vaihtoehto. Tämän jälkeen asennusohjelma päivittää San Volume Controller klusterin molemmat noodit, noodi kerrallaan. San Volume Controller ilmoittaa päivityksen statuksen GUI:ssa (Kuva 13). Vaikka San Volume Controller on päivityksen aikana toimintakuntoinen, voi palvelussa havaita suurentunutta vasteaikaa. Tästä johtuen päivitys kannattaa suorittaa levynkäytön kannalta hiljaiselle hetkelle ja siihen kannattaa varata noin 1,5 tuntia.

Name	Status	Version
node1	Offline for Upgrade	7.1.0.1 (build 79.6.1306101000)
node2	Waiting for Upgrade	7.1.0.1 (build 79.6.1306101000)

Kuva 13: Upgrade Status

4.4 Levyjärjestelmien lisääminen

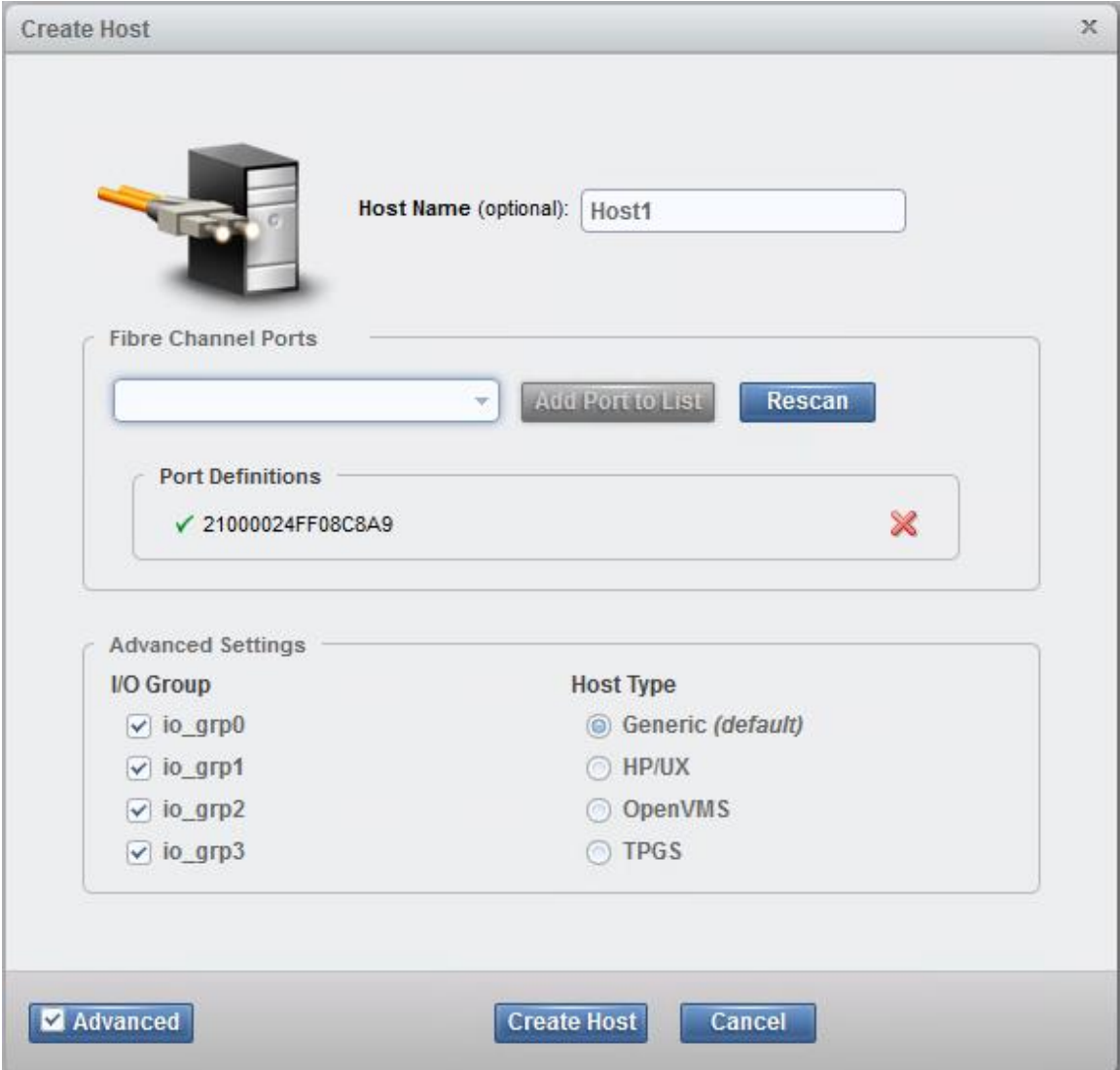
San Volume Controllerissa ei ole omaa tallennuskapasiteettia, vaan se käyttää fibre channel protokollalla liitettyjen levyjärjestelmien kapasiteettia ja esittää sen palvelimille omanaan. Ennen levyjärjestelmien lisäystä San Volume Controllerille täytyy SAN verkon zoning suorittaa ja levyjärjestelmät konfiguroida. SAN zoningin tarvitsemat WWPN:t löytyvät Settings > Networking valikosta. Levyjärjestelmiin San Volume Controller konfiguroidaan niin kuin mikä tahansa host.

Kytkinten ja levyjärjestelmien konfiguroinnin jälkeen San Volume Controller tunnistaa sille jaetut LUN:it MDiskeinä. MDiskeistä tehdään pooleja joista jaetaan palvelimille levykapasiteettia. Poolit luodaan San Volume Controllerin Pools > MDisk by pools valikosta Create Pool toiminnolla. Yhteen pooliin voidaan lisätä useita MDiskejä. Poolille annetaan nimi ja blokki koko.

4.5 Hostien lisääminen

Ennen hostien lisäämistä San Volume Controllerille täytyy suorittaa SAN verkon zoning. Tämän jälkeen voidaan San Volume Controlleriin lisätä hosteja, joille halutaan jakaa levykapasiteettia. Hostien lisääminen tapahtuu Host > Host valikosta Create Host toiminnolla. Vaihtoehtoina on Fibre Channel host ja iSCSI host. Näistä vaihtoehtoista valitaan Fibre

Channel host. Hostille annetaan nimi (Host1) ja pudotusvalikosta valitaan porttinumerot, jotka kuuluvat kyseiselle hostille (Kuva 14).



Create Host

Host Name (optional): Host1

Fibre Channel Ports

Port Definitions

21000024FF08C8A9

Advanced Settings

I/O Group

- ☒ io_grp0
- ☒ io_grp1
- ☒ io_grp2
- ☒ io_grp3

Host Type

- ☒ Generic (default)
- ☐ HP/UX
- ☐ OpenVMS
- ☐ TPGS

☒ Advanced

Create Host Cancel

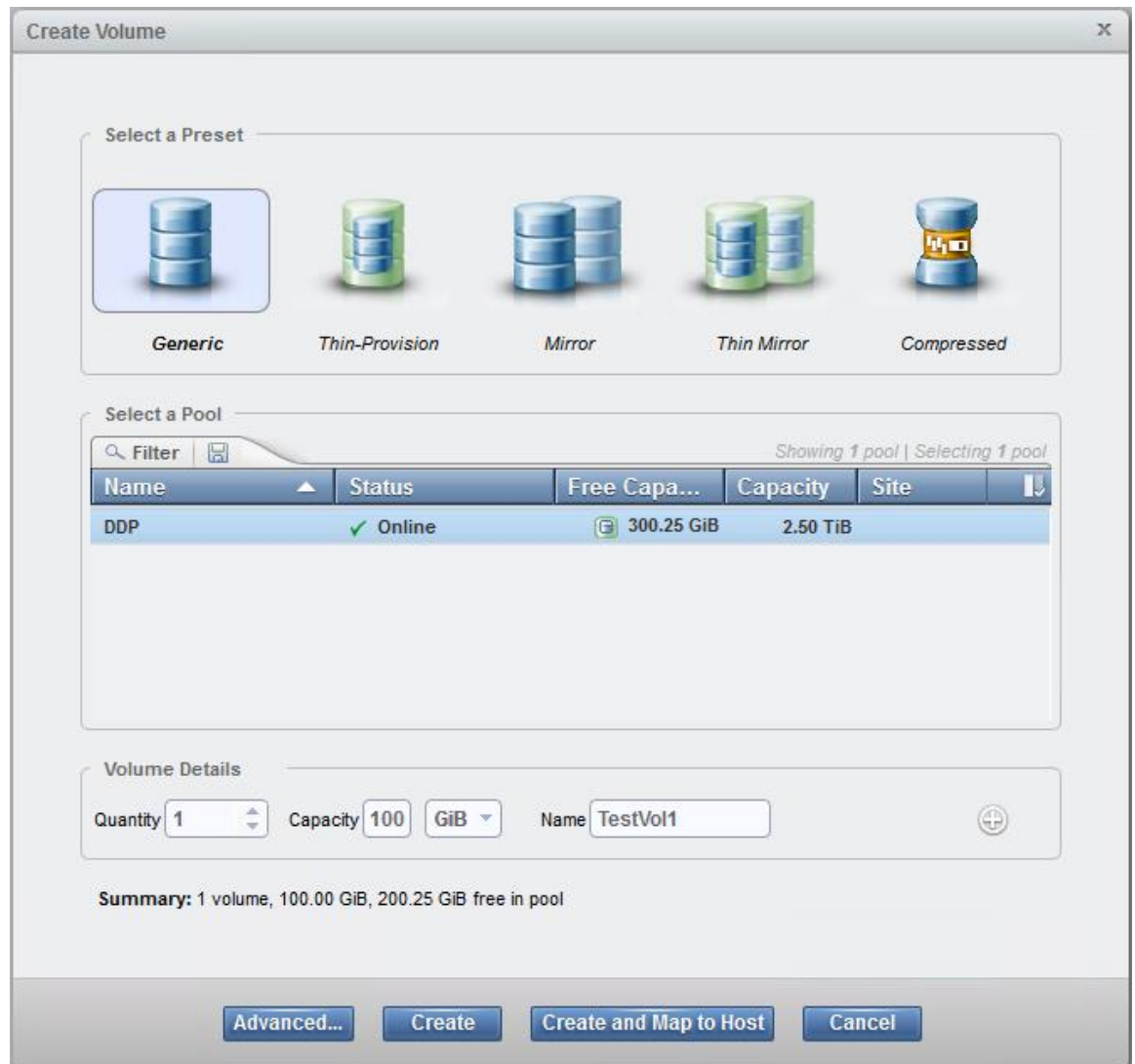
Kuva 14: Fibre Channel hostin lisääminen

4.6 Levyjakojen lisääminen

Kun San Volume Controlleriin on lisätty levyjärjestelmät mistä tallennuskapasiteetti tarjotaan ja hostit joille tallennuskapasiteettiä tarjotaan, voidaan levyjakoja lisätä. Levyjaot ovat LUNeja, joita San Volume Controllerissa kutsutaan volumeiksi.

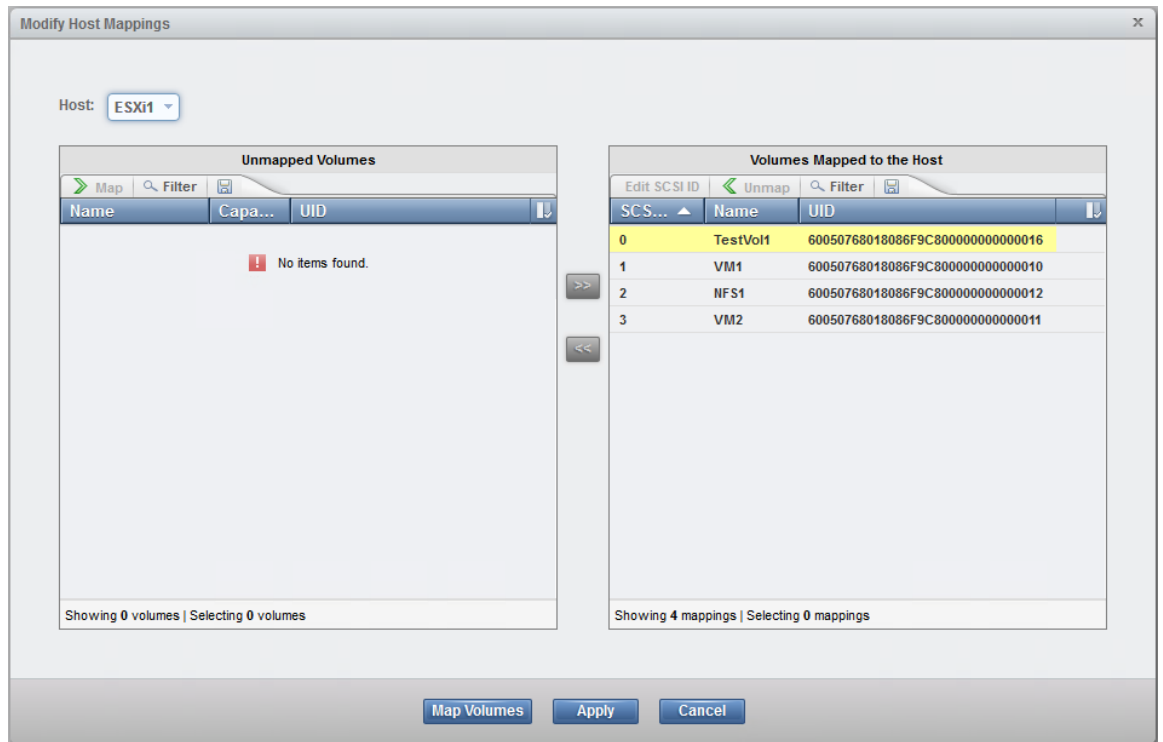
Levyjakojen lisääminen tapahtuu Volumes > Volumes valikosta create volume toiminnolla. Volumelle valitaan ominaisuudet (Normaali, thin-provisioitu, peilattu, thin provisioitu peili

tai pakattu) ja mihin pooliin se halutaan luoda. Lisäksi valitaan montako volumea luodaan, volumien kapasiteetti, sekä nimi. Jos luodaan useampi volume kerralla, käytetään juoksevaa numerointia volumen nimen lopussa.



Kuva 15: Uuden volumen luonti

Kun volume on luotu, täytyy se liittää hostiin. Tämä tapahtuu joko volumen luonnin yhteydessä Create and Map to Host painikkeella (Kuva 15) tai volumen luonnin jälkeen Volumes > Volumes valikosta valitsemalla luodun volumen asetukset. Aluksi valitaan host johon volume halutaan liittää, jonka jälkeen halutut volumet siirretään Unmapped Volumes listasta Volumes Mapped to the Host listalle (Kuva 16). Volume voidaan liittää samanaikaisesti useaan hostiin.



Kuva 16: Volumen liittäminen hostiin

5 YHTEENVETO

San Volume Controllerin fyysinen asennus on erittäin suoraviivaista. Fyysisen asennuksen suurimmaksi haasteeksi osoittautui UPS:sien kontrollikaapeleiden sijoittaminen, koska näissä kaapeleissa oli useampi metri ylimääräistä ja ne olivat erittäin paksut. Koska asennettu ympäristö koostui San Volume Controllerin ja palvelimien lisäksi vain yhdestä levyjärjestelmästä, ei mahdollisia asennus vaihtoeitoja ollut kovin montaa. Myös levykapasiteetin vähyys ja ympäristön käyttö rajoitti mahdollisia konfigurointi ja testivaihtoehtoja, eikä esimerkiksi levyjärjestelmän tilan vähyys mahdollistanut erilaisten testien suorittamista.

Mielestäni San Volume Controller on pieneen ympäristöön liian kallis tuote. Mutta jo vähänkään isommissa ympäristöissä, jossa esimerkiksi dataa halutaan peilata kahden konesalin välillä, on San Volume Controller hyödyllinen hankinta. Myös muut San Volume Controllerin ominaisuuksista ovat hyödyllisiä, kuten Easy Tier. San Volume Controllerin välimuistilla voidaan myös tehostaa hitaimpia levyjärjestelmiä huomattavasti.

Muiden virtualisointilaitteiden vertailu IBM:n San Volume Controlleriin osoittautui haasteelliseksi, koska laitteita ei ollut saatavilla ja niiden hankkiminen osoittautui liian kalliiksi.

LÄHTEET

- (1) storage virtualization Saatavilla: <http://searchstorage.techtarget.com/definition/storage-virtualization> Luettu 13.11.2014
- (2) storage hypervisor Saatavilla: <http://searchstorage.techtarget.com/definition/storage-hypervisor> Luettu 13.11.2014
- (3) LUN Saatavilla: <http://www.techterms.com/definition/lun> Luettu 13.11.2014
- (4) IOPS - I/O (Input/Output) operations per second Saatavilla: <http://www.webopedia.com/TERM/I/IOPS.html> Luettu 13.11.2014
- (5) Block level vs file level Saatavilla: <http://searchstorage.techtarget.com/answer/Block-level-vs-file-level> Luettu 13.11.2014
- (6) RAID 0, RAID 1, RAID 5, RAID 10 Explained with Diagrams Saatavilla: <http://www.thegeekstuff.com/2010/08/raid-levels-tutorial/> Luettu 13.11.2014
- (7) storage area network (SAN) Saatavilla: <http://searchstorage.techtarget.com/definition/storage-area-network-SAN> Luettu 13.11.2014
- (8) SAN vs NAS - What Is the Difference? Saatavilla: <http://compnetworking.about.com/od/networkstorage/f/san-vs-nas.htm> Luettu 13.11.2014
- (9) Kuva 4 Saatavilla: <http://www.amazon.com/QLogic-QLE2562-Adapter-Multi-Mode-Optic/dp/B00KGF2VOM> Luettu 13.11.2014
- (10) Kuva 4 Saatavilla: <http://www-03.ibm.com/systems/my/storage/san/b-type/san24b-5/> Luettu 13.11.2014
- (11) Kuva 4 Saatavilla: <http://stablefiber.manufacturer.globalsources.com/si/6008837234023/pdtl/Fiber-optic-patch/1078870934/OM3-LC-LC-Simplex-Fiber-optic-Patch-Cord.htm> Luettu 13.11.2014
- (12) Storage Protocol Comparison White Paper Saatavilla: http://www.vmware.com/files/pdf/techpaper/Storage_Protocol_Comparison.pdf Luettu 13.11.2014
- (13) vStorage - Software vs Hardware iSCSI? Saatavilla: <http://vinfrastructure.it/vdesign/vstorage-software-vs-hardware-iscsi/> Luettu 13.11.2014
- (14) Why IBM SVC and EMC VPLEX are not the same. Saatavilla: <http://www.thestoragechap.co.uk/tscblog/2012/08/05/why-ibm-svc-and-emc-vplex-are-not-the-same/> Luettu 30.12.2014
- (15) Ideas and knowledge are the lifeblood of IBM. Saatavilla: <http://www.research.ibm.com/labs/almaden/history.shtml> Luettu 13.11.2014

(16) San Volume Contoller Saatavilla: <http://www-03.ibm.com/systems/storage/software/virtualization/svc/overview.html> Luettu 13.11.2014

(17) Sangam Racherla, Matus Butora, Hartmut Lonzer, Libor Miklas Implementing the IBM System Storage SAN Volume Controller V7.2 Saatavilla: <http://www.redbooks.ibm.com/redbooks/pdfs/sg247933.pdf> Luettu 13.11.2014

(18) WHAT IS VPLEX? 2014. Saatavilla: <http://thesanguy.com/2014/02/11/what-is-vplex/> Luettu 13.11.2014

(19) Jennifer Aspesi, Oliver Shorey EMC VPlex Metro Witness Technology and High Availability Saatavilla: <http://www.emc.com/collateral/hardware/technical-documentation/h7113-vplex-architecture-deployment.pdf> Luettu 13.11.2014